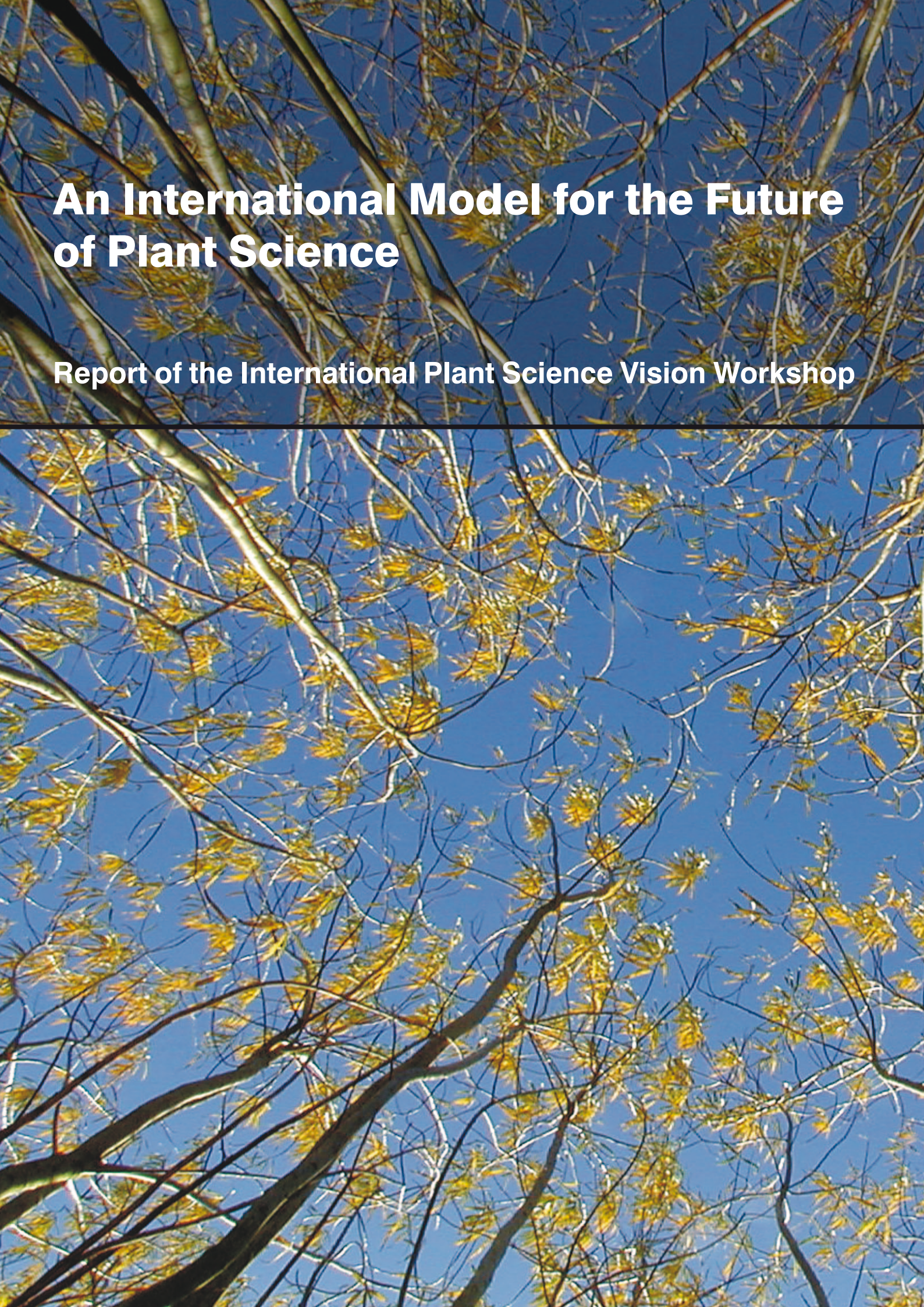
An International Model for the Future of Plant Science

An International Model for the Future of Plant Science

Report of the International Plant Science Vision Workshop



An International Model for the Future of Plant Science

Report of the International Plant Science Vision Workshop

Within the past five years, a variety of documents have been developed outlining visions for plant science with differing geographic and scientific foci (listed in section XII). This document, however, is unique in the scope and scale of its recommendations; it provides an international vision for the next decade to inform scientists, funders, and decision makers globally. Although the focus here is specifically on plant science, agronomy, soil science, the legal framework protecting plant varieties, ecology, and other areas not covered in this vision will also be crucial in translating fundamental understanding of plants into effective actions. This vision is a direct outcome of a meeting of plant science researchers from 10 different nations who were invited to the Banbury Center at Cold Spring Harbor Laboratory in the United States in 2009 to discuss their perspectives on the current and future challenges in plant science. Funding for the meeting was provided by three national funding agencies and coordinated through the International Steering Committee for Plant Genomics. The meeting participants are listed in Section X.

Table of Contents

I. Introduction

The challenges facing today's society are global. Climate change does not respect national boundaries, arable land is shrinking worldwide, and feeding the world's growing population is an ever-increasing problem. We must change the plants we grow and the way we grow them if we are to continue to meet pressing needs for food, fuel, shelter, and novel plant-based products. Meeting challenges of this magnitude will require a new way of thinking as well as the development of new approaches that are not possible with current technologies. Plant research in the next decade must reach far beyond the current focus on crops and agriculture and instead encompass all plants and plant-based products in addition to the impacts of plants on, and contributions to, the environment.

The plan presented in this vision document has been developed around the following guiding principles, which will be critical to its successful implementation:

- There is strength in diversity. The maximum benefits of plant science will only be gained by employing a variety of research systems and plant species and by drawing from the broadest range of researchers and perspectives. Plant research is carried out around the world. Its dissemination must address the needs of many nations.
- International coordination of investment and research efforts will be essential to leverage national investments and synergize outcomes that lead to significant change. The grand challenges in plant biology are larger in scale than is feasible for any single agency to support, and translation of research into new products will require new kinds of cross-disciplinary approaches that will likely transcend national boundaries.
- A key output from plant science in the coming decade will be a universally accessible, modular tool kit that will enable rapid responses to unforeseen challenges, such as the emergence of new plant diseases, changes in the distributions of plant pathogens and vectors, and the emergence of new weather patterns.
- Timely and open access to all of the tools, resources, and data generated through internationally coordinated projects will be essential to the success of this vision. Without unfettered access, the benefits of the investments in plant research will be reduced or lost.

The past two decades have brought increased investment in plant science coupled with an expansion of international coordination that was brought about, in part, through the efforts of national funding agencies as well as multinational flagship projects in plant genome sequencing and functional genomics. These successes have set the stage for future efforts in plant science at an international level. This document describes a bold, decadal vision for plant science that will maximize the impact of research investments to develop outcomes in the field with benefit to all.

II. Executive Summary

Plant biology is needed more today than ever before to meet the demands of a growing global population and the challenges of global climate change while safeguarding the environment. A leap in global plant productivity (of biomass, food, and fiber) is now required and can only be secured in the pressing timeframe through a quantitative understanding of plant biology, including the plant's interaction with soils and the effects of human interventions, and translating these findings efficiently to crop production and the development of plant-based products.

A revolution driven by genome-scale technologies, remote sensing, systems biology, and informatics is transforming the practice of plant research, which is our focus here. Plant science must therefore balance three goals:

- **Understanding** the fundamental processes of plant biology with sufficient accuracy to build a mathematical model of the plant from its molecular parts;
- **Deciphering** the particular biological processes that underlie economically important plant traits;
- **Applying** the new understanding and models to develop new and improved crops and to protect the environment for the decades ahead;

These goals present the research community with challenges at four levels that demand new research tools and an internationally coordinated infrastructure:

- New experimental methods are required to locate and quantify all molecular components of plant cells, from small chemicals to complex structures, so that genome sequences can be linked to physiological outcomes. Arabidopsis will remain the reference plant species for these studies, but a small number of additional species should also be adopted for focused study. Genome sequence data should be obtained for a wider diversity of plant life along with high-throughput measures of plant performance in diverse conditions. A global policy on access to genetic resources will be required to gain the best value from the wealth of national and international germplasm banks.
- The integration of current "'omics" data, such as genomics, transcriptomics, proteomics, metabolomics, interactomics is not a solved problem, but the decadal goals of plant science require even more than this from our **data infrastructure**. Quantitative data on molecular components must now be linked to mathematical models and geographical maps, potentially in "real" time as a plant grows or an epidemic unfolds. Although the plant science community to date has successfully provided shared data resources, open and timely access to pre-competitive data and interpretation will require a new commitment to coordination and funding and may require community-approved enforcement.
- Formal (mathematical or computational) models offer the only viable means to understand how plant performance through the entire life cycle -"from seed to seed" - derives from interactions among the molecular parts. New methods will be required for modular **modeling across multiple scales** to link genome sequence and quantitative data to the mechanisms of plant growth and development as well as to outcomes in economically and scientifically important traits or in ecosystems.
- An effective knowledge exchange **pipeline** must link laboratory testing to validation in field conditions for many species. Establishing this pipeline is a genuinely innovative project covering human resources, interaction across public and private sectors, and international regulatory reform. A wholesale reorientation of the educational programs in biological sciences will be required to multiply the numbers of interdisciplinary scientists and plant breeders adept in using and developing new technologies.

There must be substantial and sustained multinational support to achieve these goals, reversing the trend of reduced research funding in this area (Alston *et al.*, 2009). Our vision is to build from the first isolated examples of technology-enhanced crop improvement to a systematic process that can be implemented on an unprecedented scale in plant species that range from traditional row crops and wild plants to trees and even algae. This vision depends on deeper international cooperation not only to tackle the "Grand Challenges" in research and to adapt to diverse needs, but also because projects of high international relevance and urgency cannot wait for funding and organization to emerge piecemeal from local initiatives.

III. Societal and Scientific Justification

The food price surges and food riots that sprang up around the world in 2008 were a stark warning that feeding an additional 2.2 billion people by 2050 (Figure 1) will not be an easy task. However, this is one area where plant science will make a decisive contribution. The doubling of world food production over the past half century has now slowed (Figure 2) (*Alston et al*., 2009) and came at a cost. Much of the best agricultural land is already in use (*Goldweijk et al*., 2007), and a significant fraction is being lost due to overtillage, overuse of fertilizers, salinization, erosion, urbanization, and industrial contamination (Scher and Yadav, 1996). The genetic diversity of plant varieties in production has narrowed, and biodiversity in natural systems is under threat. In the next 40 years, food and feed production will need to double again while restoring and maintaining the quality of the planet's natural resources for future generations. Failure to achieve these goals could have devastating consequences.

The use of plant genetics was responsible for half of the historical increase in crop yields. Future success will similarly depend on the development of novel, high-performing crops. In the coming decades, we will face new challenges, including severely reduced water availability, global warming and the development of volatile weather patterns (Burke *et al*., 2009), competition from non-food crops (Campbell *et al*., 2009), and urban encroachment as the growing population moves from the farms to the cities (Doos, 2002). New crops must therefore combine higher yield with improved quality; with tolerance to drought, temperature, and other stresses; and with resistance to pests and disease. Better agricultural practices, improved national infrastructures, and imaginative new international policies will also be needed to ensure that a global population of nine billion people can be fed equitably and sustainably, but increased and stable crop yields are the crucial foundation.

Plant research is at the foundation of all of biology. A consistent feature of this research is the crucial role played by curiosity-based research, that is, research undertaken by plant scientists to satisfy their own curiosity about the underlying processes of plant life. It is such plant research that led to the discoveries of cells in the 17th century; of nuclei, viruses, and the laws of genetics in the 19th; of transposable elements and the mechanisms that capture solar energy and CO_2 into organic matter in the 20th; and of small RNA-based gene regulation in the 21st. In addition, the functions of conserved genes in protein degradation, photoperception, and many other areas were first determined in plant systems. Undoubtedly the unpredictable discoveries yet to come in plant science will bring fundamental, new dimensions to our understanding of biology. Investigator-initiated, fundamental research has proven critical to providing such discoveries and the novel technologies they require.

Plant science is now at the beginning of a revolution—driven by genome-scale and imaging technologies, systems biology, and informatics—that will underpin the next generations of economically important plants. The research that will contribute to this revolution spans many areas, including the mechanisms of development and growth, regulatory and metabolic networks, and the molecular basis of evolutionary and adaptive change. A major challenge will be to achieve these fundamental advances while ensuring that the results flow rapidly to practical plant breeding across many species. Combining research on model species with new approaches to breeding and genetic modification offers a demonstrated acceleration in plant improvement. New paradigms will be required to develop plants for bioproducts, including biofuels and pharmaceuticals, while preserving biomes and biodiversity. Entirely new crop species will be developed and evaluated in this direction.

The next decade will be critical for plant science. Increased food production will be required to sustain the dramatic population jump expected within the decade. The policies, the investment, and the science conducted in the coming years will determine how successfully we meet the greatest and most immediate challenge ever faced by the human race.

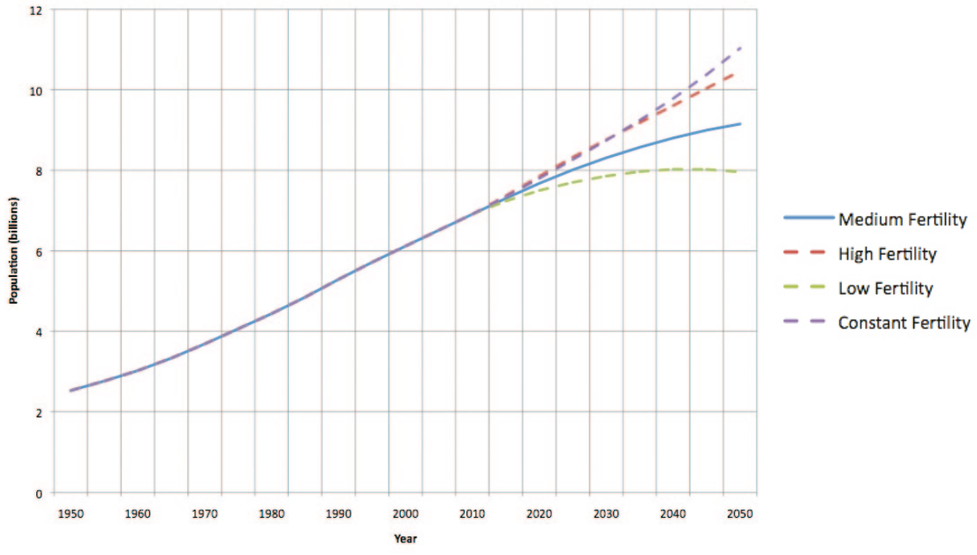


Figure 1: Graph of projected world population prospects, 1950-2050. *Source - Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, World Population Prospects: The 2008 Revision http://esa.un.org/unpp/*

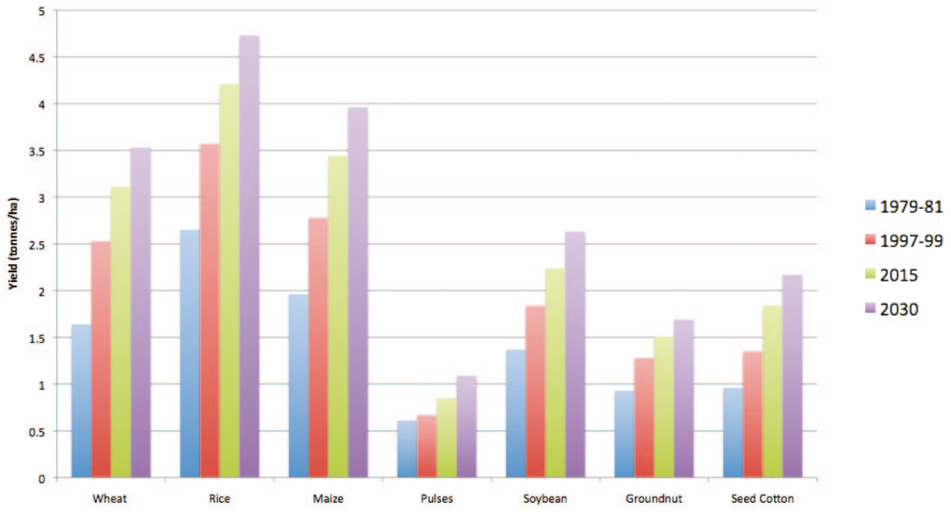


Figure 2: Crop Yields in Developing Countries 1979-2030. *Source - FAO Economic and Social Development Department: World Agriculture towards 2015/2030 Report http://www.fao.org/*

IV. Decadal Goals in Plant Science

Research in plants should target four transformative goals:

(I) To understand the central processes of plant biology with sufficient accuracy that a mathematical model of the plant can be built from its molecular parts;

(II) To decipher the particular plant processes that underpin agriculture and other plant-based industries;

(III) To integrate this new understanding of plant biology with a model-based approach towards the design of new plants yielding food, fuel, and biomaterials for the decades ahead; and

(IV) To evaluate plant biodiversity at a global scale to unravel the unique biological processes that have major importance to environmental response and adaptation.

In understanding the central processes of plant biology, plant science should aim to:

- Gain a molecular-level understanding of how structural, regulatory, and metabolic networks, including epigenetic processes, are integrated in living plant cells;
- Understand how cells constrained within cell walls can communicate to form an entire organism and how they interact with microbes;
- Deconstruct the mechanisms that plants use to integrate the internal, developmental, and environmental signals that control their responses in the ecosystem;
- Understand the molecular basis of evolutionary change in key species over longer timescales and how this links to biodiversity in mixed communities.

In uncovering the biological foundations of agriculture, plant science should aim to:

- Decipher the biological processes underlying key plant traits including heterosis, nutrient use efficiency, drought tolerance, apomixis, perenniality, self-incompatibility, and the control of recombination;
- Understand the evolution and genome dynamics involved in the domestication of crops from wild plants;

To apply new understanding and develop dynamic models, plant science should aim to:

- Link the knowledge gained from understanding of these regulatory networks and evolutionary driving forces to create plant ideotypes—idealized plant types with specific combinations of favorable characteristics;
- Move plant breeding from empirical approaches to a predictive science using models of plant growth and behavior based on molecular data—the "computable plant";
- Preserve the environment and its interrelationships with agriculture and other plant-based industries by applying knowledge of the molecular basis of biodiversity.

Realization of these diverse goals will allow the decoding of plant life through the entire life cycle, "from seed to seed," and translation of this knowledge into improved plants with enhanced agronomic properties.

V. Challenges

Plant cells process many inputs from their environment and from the plant itself to achieve adaptive responses in metabolism and development. Each plant's particular combination of adaptive programs must be decoded in terms of specific, underlying biological processes. In parallel, our understanding of each molecular process must be integrated in order to reassemble models of a whole plant. Interactions with other species and geophysical constraints must be captured to understand how plants operate within ecosystems. Further, the agroeconomy must be reconciled with the need to preserve biomes and germplasm collections, valuing them as essential sources of plant diversity. Within the last decade, an impressive amount of data on these processes has already been obtained, and we are now confronted with a new set of conceptual challenges.

A. Development of multidisciplinary approaches

Comprehensive models of plant growth offer the only viable means to integrate a wide range of information and apply it to address societal and environmental challenges. There is widespread agreement that new types of mathematical modeling are required to understand plant biology in an integrative manner. Detailed models of particular subsystems should be combined within a larger theoretical framework to understand how plant performance from seed to seed derives from interactions among the parts, interactions between the parts and the plant's environment, and how individual plants contribute to mixed communities. This approach requires new theoretical concepts for modular modelling across multiple scales in time and space, including multiple physical

Figure 3: Wild type strains of the model plant *Arabidopsis thaliana*
As a result of its small size, rapid generation time, small genome and ability to generate a vast amount of seed by self-pollination, Arabidopsis has emerged as the organism of choice to study plant biology in laboratories across the globe over the past 20 years.
Image kindly provided by Janne Lempe, Sureshkumar Balasubramanian and Detlef Weigel

processes (diffusion, mechanics). Such methods will be equally applicable to modeling other multicellular organisms.

B. Extension of studies to a larger number of reference systems

Arabidopsis (Figure 3) will remain the foundational reference plant species, partly for practical reasons of efficiency in the introduction and validation of new approaches, partly because of the availability of focused technologies and tools, and partly for conceptual reasons arising from the depth and concentration of biological understanding. However, plant science will move beyond this pioneering model species in order to translate current understanding to economically important plant species, to enable comparative studies that enhance understanding of all plants, and to study traits that are otherwise inaccessible. The adoption of a small number of additional reference species from unicellular algae to trees will expose more complex genomes and biology to detailed study. Sequencing technology will rapidly make available genome data for a still wider diversity of plant life, including "orphan," or minor, crop species and species with unusual adaptations.

C. Coordination of data management across the community

Simultaneous transformations in the rate and diversity of data acquisition, together with the wider range of species under study, risk overwhelming the research community's ability to assimilate the experimental results. The conceptual challenges in data integration and data management will limit our ability to extract scientific value, especially at the network-wide scale. The success of predictive biology and the dissemination of fundamental advances to many species across the globe depend upon timely, open, and informed access to research results. The data integration issue must be addressed in a coordinated manner across the international plant research community, recognizing linked issues such as attribution and intellectual property.

D. Translation of fundamental plant science to applications

The models derived from fundamental research must extend beyond the laboratory to the complexity of natural and field environments. In order to create plants capable of producing dependable, robust yields in the face of changing environmental conditions, an effective knowledge exchange process must link genome assembly and testing in the laboratory to phenotyping and validation in field conditions (Figure 4). Systems biology models can thus be linked to the ongoing modeling traditions of crop science and ecology, and metagenomic sequencing can inform research on biodiversity and climate change. The establishment of a pipeline from the laboratory to the field is a genuinely innovative project that will require a multi-faceted approach, covering human resources, interaction across public and private sectors, and international regulatory reform.

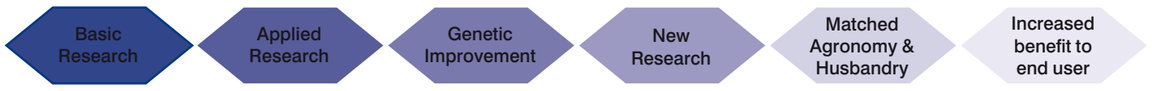


Figure 4: Pipeline of plant research from the laboratory to the field

Plants are complex systems, that is, they are composed of elements with particular properties that interact to form structures at a higher scale with emergent properties depending on these interactions. Thus, molecules assemble into complexes, which assemble into cells, tissues, organs, and finally entire plants. Plants, in turn will interact with their environment as part of ecosystems. If we want to understand how these assemblies function, we need to study the components at each level of organization. More particularly, we must define what the parts are, how they interact, and the dynamics of their interactions. In this context, quantitative approaches will be central at all levels of organization, but these raise particular requirements:

A. Intracellular level

A complete inventory is now required of all the molecules that make up plant cells, from enzymes to metabolites and including their spatio-temporal distribution within the cell or plant. This information must be obtained using quantitative biochemical assays, ideally reaching single-cell and single-molecule resolution, including proteomic and metabolomic approaches.

In vivo imaging and spectroscopic techniques will be essential to locating and quantifying the molecular components of plant cells. Large collections of protein-tagged transgenic plants can be developed rapidly as a first step to localize individual proteins. However, detection methods must be extended to locate and quantify small molecules, such as hormones, sugars and micronutrients.

The next challenge will be to identify how these parts interact to form a robust molecular network. Mapping and modeling the genetic regulatory network has been started using a range of "'omics" approaches focused on transcriptional regulation. Essential quantitative and spatial information is still missing regarding transcription, epigenetic regulation, protein stability, interactions with metabolism and dynamic distribution. This must now be provided using new experimental techniques. However, the regulatory network cannot be described by the activity of genes and transcription alone. It also depends on other cellular components, for example, when cytoskeletal elements transport regulatory molecules. Efforts should therefore be aimed at elucidating how the chemical constituents interact to form physical structures, such as how the cell wall assembles.

Box 1. 'Omics' Approaches

'Omics' is a general term for the broad discipline within science of the analysis of biological information objects in various 'omes' including the genome, proteome, and metabolome. 'Omics' approaches can be thought of as the tools and technologies that allow the analysis of this biological information often at large, high throughput scales. 'Omics' approaches include:

Genomics - study of gene sequences or genomes

Transcriptomics – study of the expression levels of (m)RNAs

Proteomics – study of a large-scale set of proteins

Metabolomics – study of a large-scale set of metaboltites

Interactomics – study of genome scale interactions amongst proteins

Epigenomics – the study of the chemical modifications and structural features such as histone modifications and their distribution across the genome

The structural, metabolic, and regulatory molecular networks can only be understood using integrative approaches that combine diverse data types into models that represent spatial structure in detail. Ultimately, this should enable us to understand how a physical, dynamical structure such as the plant cell is assembled, how it responds, and how it functions.

B. Cellular and intercellular levels

Cells interact biochemically and physically to form tissues in which differentiated cell types are characterized by distinct functions (photosynthesis, uptake of nutrients, development of reproductive structures) and, importantly, distinct growth dynamics.

In the last decade, considerable advances have been made regarding major signaling pathways that coordinate cell behavior. Examples include the light regulation of gene expression and hormone signaling pathways. By contrast, we do not know the functions of hundreds of receptor-like molecules and putative peptide hormones that have been identified using genomics. The potential cell-to-cell signaling functions of metabolites and RNAs remain largely unknown. Cells also interact physically. The physical interactions are translated into cellular responses via structural components, such as the cell wall and the cytoskeleton. New methods will be required to elucidate these little-known processes.

Molecular regulatory networks interact with cells' structural elements to determine the cells' size and shape through modification of the cell wall. This is a major issue in developmental biology that is not understood in detail for any cell. Quantitative understanding will require comprehensive analysis of growth patterns with cellular resolution combined with dynamic perturbations of the regulatory network.

C. Organismal level

The molecular networks that make up cells and tissues, in interaction with the environment, ultimately lead to the formation of a plant with a specific size and architecture. As the behavior of the whole plant cannot be predicted from the properties of molecules and cells alone, we must continue to study the plant at the scale of tissues and organs. Quantitative measurement of growth patterns in a range of environments has led to the development of descriptive models. Integrating molecular and cellular information will greatly enhance the models' predictive power. Such multiscale models will link genotype to phenotype, for example, connecting genomic information arising from next-generation sequencing with specific plant traits at the whole-plant level. Connecting environmental inputs to whole-plant responses will require precision phenotyping methods that can be applied in the field and in natural environments.

The analysis above is applicable to all plant species and refers to common mechanisms. Comparative analyses will highlight the differences among species that lead to their distinct functional characteristics and the relevant mechanisms of genome evolution. Plants offer particular opportunities to unravel the evolutionary consequences and opportunities arising from hybridization, polyploidy, and genome duplications.

D. Supra-organismal level

The plant's place in the environment exposes it to abiotic factors as well as to other organisms, both beneficial and harmful. Metagenomic analysis of the phytosphere, investigating phytopathology and symbiosis, will resolve genetic diversity in unprecedented detail, potentially tracking evolution in near-real-time. Remote sensing methods will monitor plant performance in many locations, also potentially in real time. Combined understanding of genetic diversity and environmental responses will reveal the spectrum of selective pressures and the molecular origins of adaptive change. In the context of crops in the field, this integrated understanding of genotype, phenotype, and agronomic environment is the foundation for the specification of crop ideotypes that match local conditions. Actually realising these ideotypes will depend on a greater understanding of the processes that organise plant genomes (such as recombination), and will demand enhanced transformation technologies for rapid and specific manipulation of plant genomes.

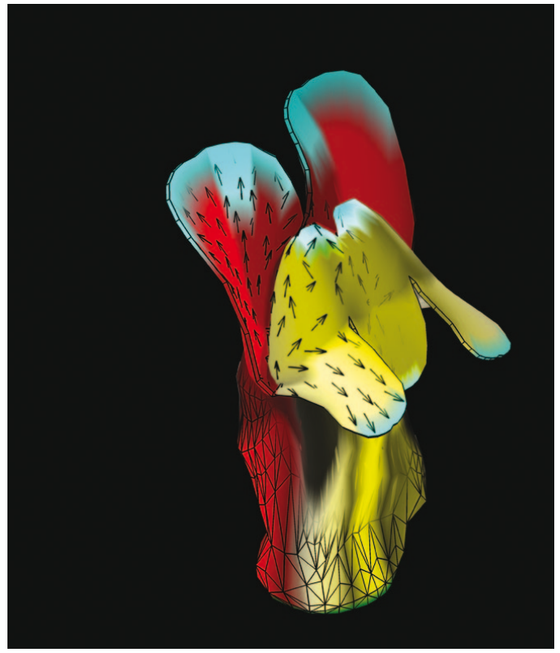


Figure 5: Computational model of a snapdragon flower
Source - Green AA, Kennaway JR, Hanna AI, Bangham JA, Coen E, 2010 Genetic Control of Organ Shape and Tissue Polarity. PLoS Biol 8(11)

E. Common strategies across levels of organization

The data regarding the interactions of hundreds of components (molecules, cells, or plants) in time and space will be impossible to grasp without the appropriate informatics and modeling tools. Integrated databases are needed to connect different data types and accelerate the development of quantitative models. Models in the form of virtual cells and tissues are being developed (Figure 5), but complex spatial representation remains a challenge. Multiscale modeling will require new theoretical frameworks that populate the spatial structures with biochemistry, physics, and cell biology. This area is already testing computing power to the limit and will require innovative software and hardware. Both theoretical and technical developments will then be required to apply the systems approach on a large scale: first to build the models and then to automate data acquisition, analysis, and model parameterization, in order to test the effects of many environmental signals, genetic manipulations, or chemical perturbagens.

VII. A Global Infrastructure for Plant Science

The next five years will see a revolution in the amount of genome-scale data available for plants. Not only will the genome sequence of all important crop and model species be deciphered, but also so will the sequences of multiple varieties of these species, unlocking the wealth of collections of germplasm amassed over the past 150 years. This data will be supplemented by extensive sampling of genetic variation in wider populations, equivalent data from plant pathogens and symbionts, and parallel studies of gene expression and regulation. As new technologies revolutionize both the volume of available data and our potential to address ever-larger challenges, the importance of developing appropriate, internationally coordinated infrastructure and human resources is growing. The plant science community has demonstrated profound commitment to and successful provision of shared resources for data acquisition, dissemination, and analysis. However, research resources differ among nations, requiring flexible mechanisms to provide appropriate engagement. Three essential pillars will help to build upon this foundation and to accelerate international research in plant science over the next decade:

Figure 6: Wealth of Plant Genome Data
Table of plant species, which either have a published genome sequence or that are currently in the process of being sequenced.
This is information has kindly been provided by James Schnable http://synteny.cnr.berkeley.edu/wiki/index.php/Sequenced_plant_genomes

Definitions
Published: *A complete genome sequence is available, and anyone can publish papers on it without restriction.*
Unpublished: *The complete sequence (or a substantially complete sequence) is available, but whole genome analysis cannot be published until the group that sequenced the genome publishes their own paper describing it. These restrictions are outlined by the Fort Lauderdale Convention.*
Incomplete*: A partial assembly is available, but sequencing and/or assembly and/or gene annotation is ongoing*
Unreleased: *Genome sequencing has at least begun, but no data has been made publicly available*

Organism	Common Name	Published	Unpublished	Incomplete	Unreleased
Aquilegia coerulea	Colorado blue columbine		✓		
Arabidopsis arenosa	Sand rock cress				✓
Arabidopsis halleri	Rock cress				✓
Arabidopsis lyrata	Lyre-leaved rock cress		✓		
Arabidopsis thaliana	Thale cress	✓			
Arachis hypogaea	Peanut				✓
Beta vulgaris	Sugar beet				✓
Brachypodium distachyon	Purple false brome	✓			
Brassica napus	Oilseed rape				✓
Brassica rapa	Chinese Cabbage				✓
Brassica oleracea	Various				✓
Boechera divericarpa	Purple rock cress				✓
Boechera holboellii	Holboll's rock cress				✓
Boechera stricta	Drummond's rockcress				✓
Capsella rubella	Pink Shepherd's Purse				✓
Carica papaya	Papaya	✓			
Castanea mollissima	Chinese Chestnut				✓
Citrullus lanatus	Watermelon				✓
Citrus clementina	Sweet orange		✓		
Citrus sinensis	Clementine mandarin		✓		
Coffee caneophora	Coffee				✓
Cucumis sativus	Cucumber	✓			
Eucalyptus grandis	Rose Gum Tree		✓		
Fragaria vesca	Woodland strawberry	✓			
Glycine max	Soybean	✓			
Gossypium raimondii	Cotton				✓
Hordeum vulgare	Barley				✓
Lactuca sativa	Lettece				✓
Malus x domestica	Apple	✓			
Manihot esculenta	Cassava			✓	
Medicago truncatula	Barrel medic		✓		
Miscanthus giganteus	Giant silver grass				✓
Mimulus guttatus	Monkey flower			✓	
Musa acuminata malaccensis	Banana				✓
Oryza glaberrima	African cultivated rice				✓
Oryza sativa	Rice	✓			
Panicum hallii	Hall's panicgrass				✓
Panicum virgatum	Switchgrass				✓
Phaseolus vulgaris	Common Bean				✓
Phoenix dactylifera	Date Palm			✓	
Phycomitrella patens	Moss	✓			
Pinus taeda	Lobolly Pine				✓
Populus nigra	European popular				✓
Populus tremula	European Aspen				✓
Populus trichocarpa	Poplar	✓			
Prunus avium	Sweet cherry				✓
Prunus domestica	European plum				✓
Prunus persica	Peach		✓		
Ricinus communis	Castor bean	✓			
Rubus idaeus cv. Heritage	Red Raspberry				✓
Salix purpure	Purple willow				✓
Selaginella moellendorffii	Spikemoss		✓		

Organism	Common Name	Published	Unpublished	Incomplete	Unreleased
Setaria italica	Foxtail millet		✓		
Sorghum bicolor	Sweet Sorghum	✓			
Solanum lycopersicum	Tomato			✓	
Solanum tuberosum	Potato			✓	
Spirodela polyrhiza	Duck Weed				✓
Thellungiella halophila	Salt Cress				✓
Thellungiella parvula					✓
Theobroma cacao	Cacao	✓			
Triticum aestivum	Wheat				✓
Vitis vinifera	Grape	✓			
Vitis vinifera var Sangiovese	Sangiovese grapevine				✓
Zea mays	Maize	✓			
Zostera marina	Sea grass				✓

A. Reference resources for biodiversity

National germplasm banks together with those of international centers hold an underutilized wealth of plant diversity. Technological advances offer the possibility of identifying valuable genes and gene combinations at a genome-wide scale. A rational, global genetic resources policy will be essential to ensure that this work is coordinated and that the results are available to all, thereby gaining best value for research and plant improvement.

B. An informatics infrastructure to support predictive biology

Ultra-high throughput sequencing technologies will provide extensive evidence for gene expression and regulation; the availability of more reference genomes will increase our ability to infer function in species less directly studied. Remote sensing will contribute streams of plant phenotype information from field locations (Figure 7). Yet even the relatively simple integration of 'omics data' is not a solved problem, with both informatics and organizational factors promoting data fragmentation and redundancy. If the potential offered by high-throughput data generation is not to be wasted, and if predictive biology is to be realized, we will require the ability to link such data to quantitative chemistry, developmental models, and geographic information systems and to provide systems offering access to such data in a real time frame (for example, that defined by the growth of an organ or the spread of an epidemic).

We need an infrastructure that will facilitate both the development of leading-edge modeling techniques and their subsequent application to real world problems. This should be constructed using a layered approach, building plant-centric services on top of universal resources shared with other domains of biology.

Essential elements within this system include:

- Compulsory use of public archives and established data standards for primary data and metadata (such as, for DNA sequences, those maintained by the National Center for Biotechnology Information (NCBI), the European Bioinformatics Institute (EBI), and the DNA Data Bank of Japan (DDBJ)) as enforced by funders, publishers, and the wider community. The development of an open informatics platform and public database for plant images and phenotyping is a priority.
- Use of standard informatics platforms for integration of 'omics data', to avoid unnecessary duplicative development and to provide a unified interface for downstream researchers.
- Use of standardized vocabulary and a protocol for data sharing and integration.
- Community-led resource development for individual species and groups of related species to define reference sequences, genes, etc., and to take charge of data integration and curation, according to the priorities of each community.
- High-level resources to provide an integrative, comparative view across the plant/plant pathogen/plant symbiont domains.
- Links between stock collections and 'omics-centric' databases that are seamless and efficient.
- The informatics framework needed to support predictive systems biology is not yet fully specified, and its definition will require contributions from scientists trained in many disciplines outside plant science, including mathematics, physics, chemistry, computer science, and ecology. Special initiatives will be needed to deliver the services that will nurture this inter-disciplinary research.

C. Coordinated international activity for strategic priorities and grand challenges

Projects of high international relevance and urgency cannot wait for funding and organization to emerge piecemeal from local initiatives. Such projects range from the implementational—for example, the determination of reference genome sequences for the most important plant species or the establishment of a coordinated informatics infrastructure for 'omics-scale data'—to conceptual "grand challenges" aimed at overcoming the key bottlenecks to the next generation of plant science, such as the development of multiscale modeling frameworks for systems biology. Coordination is needed and will require the development of new mechanisms to support the international harmonization of research goals and access to funds.

International policies will be required to regulate issues of ethics and IP for both biological materials and data infrastructure, but their details are beyond the scope of this document. The seminal agreements already reached in particular domains indicate the value of inclusive processes for the dissemination of innovative concepts in this area.

Figure 7: Remote sensing equipment *'Phenomobile Mach 2' and 'The Blimp' used at the high resolution plant phenomics centre, Australia to collect field measurements such as canopy temperature, canopy volume/biomass, ground cover and aerial images.* ***Source - Image kindly provided by Carl Davies, CSIRO Plant Industry.***

VIII. Foreseeable Outcomes

The leap now required in global plant productivity (biomass, food, and fiber) can only be secured in the pressing timeframe through quantitative understanding of fundamental mechanisms in plant biology and by translating these findings efficiently to innovation in bio-based industries. Concrete examples from contemporary crop improvement programs have already demonstrated the relevance of the "model species + genetic modication" approach, and the rapid progress that can been achieved, relative to conventional breeding approaches. Our vision is to build from these first, isolated examples to a systematic process that can be implemented on an unprecedented scale in plant species from traditional row crops to novel crops, to trees and even algae. Innovative agronomic practices and other inputs will unquestionably contribute to the value of this approach.

A new framework will allow the multiscale modeling of regulatory and metabolic networks within, the development of, and environmental signals found in multi-cellular organisms. Initially, these models will predict the impact of plant genotypes on responses to environmental challenges for particular responses. At a later stage, more comprehensive models will specify optimal ideotypes for sustainable agriculture.

Sequence-based measures of genetic diversity combined with phenotypic analyses will enhance the effective management of global genetic resources enabling the preservation and capture of biodiversity. This diversity will be a source of valuable genes, which may be used in developing crop varieties and plant-derived bioproducts.

Transformation technology combined with marker-assisted selection and molecular understanding of gene networks involved in conferring specific traits will allow **the rapid development of new crop varieties with high yield and yield stability**. Varieties engineered to contain genes responsible for desirable traits, such as efficient nutrient and water use, tolerance to environmental stress, and disease resistance, will support sustainable agricultural systems. It will be possible to grow these newly-bred or -engineered crop varieties in marginal lands, such as water-limited, water-logged, or high salinity areas, increasing total crop production without the need to create new farmland, thereby allowing the preservation of forests in particular. In addition, the new varieties for a wide range of crops, including those important in developing countries, will maintain production under changing climates.

Synthetic metabolic engineering with genes from plants or other organisms will increase the range and quantity of **bioproducts**—for example, pharmaceuticals may be created by engineering novel genes in biochemical pathways—and for increased production of valuable existing compounds.

From the knowledge of the gene networks involved in the domestication of existing crops, the conversion of **wild species to high value crops** will be accelerated. This can offer, for example, new alternatives to fossil fuels that need not compete with food production for land.

Synergies from **international coordination** of research and exchange of practices, genetic resources, databases, tools, and personnel will be captured for the international research community. Worldwide data integration will accelerate the attainment of future research goals and reduce duplicative activities. This organization of knowledge constitutes a new type of global intellectual capital that will lay the foundation for progress in the life sciences in future decades. For example, we anticipate complete plant genome synthesis based on ideotypes, designed at nucleotide resolution, that reliably deliver desired plant characteristics.

IX. Recommendations

A. Funding science and infrastructure

Funding organizations need to provide major and sustained support, approximating the scale of biomedical research, to capture and integrate molecular, cellular, organismal, and ecological data, to understand how living plants develop, function, and adapt to their environment, and to conduct practical crop improvement and development in the light of this knowledge. The experimental focus should be on dynamic and quantitative data at multiple scales. In parallel, new levels of effort and investment will be needed to secure interdisciplinary research that delivers modeling and data integration systems. The new plant science should be complemented with chemical biology, agronomy, land use, and allied research, though these areas are outside the focus of this report.

B. Data capture and open access

All generated data needs to have a funded path to integration and complete downstream dissemination over the period of its usefulness. Open and timely access to pre-competitive data and interpretation provides a strategic advantage to the community, which is currently not being realized and may require community-approved enforcement mechanisms.

C. Training and education

A wholesale reorientation of the educational programs in biological sciences will be required in order to multiply the numbers of interdisciplinary scientists and of plant breeders adept in new technologies. Plant science needs to exploit its competitive advantages (culture of data accessibility, international outlook, success in practical application, and moral imperatives in ecological and societal goals) in order to attract expertise in theoretical disciplines. Programs should be introduced to promote direct interaction between laboratory researchers and plant breeders in order to gain value from combining new ideas and technologies with the existing skills of the breeders. Extension to training practitioners will be essential to any widespread implementation.

D. International coordination in research and access standards

1. Research coordination

Plant science has benefited tremendously from collaborative international funding mechanisms. The vision for a coordinated, international effort towards whole-plant models crucially depends upon deeper cooperation, with international agreement of scientific priorities and strategic commitment to international implementation and delivery mechanisms. Joint funding programs should be broadened to encompass additional nations and disciplinary areas, with a substantial increase in funding, in order to gain the most value from unique infrastructures and scarce human resources. The needs of developing countries must be explicitly addressed. Support for exchanges of researchers at all career stages should be increased to broaden their direct experience, extend best practices, and maximize the use of specialized research facilities.

2. Streamlining the regulatory framework

Standardizing and simplifying regulatory processes at all levels will be required to facilitate open access to data, resources, and materials. Plant science should adopt international standards in this area, where they are agreed by the relevant community. The success of translational research from the laboratory to the field has been hampered by intractable, non-science-based regulatory processes. International dialogue should be accelerated to simplify and rationalize this process.

X. Contributors

Participants

Liz Dennis, CSIRO, Australia

Joe Ecker, Salk Institute, USA (Co-Chair)

Mike Gale, CGIAR

Klaus Harter, University of Tubingen, Germany

Paul Kersey, EBI, UK

Suk-ha Lee, Seoul National University, South Korea

Minami Matsui, RIKEN Plant Science Centre, Japan

Andrew Millar, University of Edinburgh, UK (Co-Chair)

Isobel Parkin, Agriculture Canada, Canada

Marie-Anne van Sluys, Universidade de São Paulo, Brazil

Jan Traas, INRA/CNRS/ENS, Lyon, France

Observers

Ruth Bastow GARNet, UK

Catherine Kistner, Deutsche Forschungsgemeinschaft (DFG), Germany

Rowan McKibbin, Biotechnology and Biological Science Research Council (BBSRC), UK

Jane Silverthorne, National Science Foundation (NSF), USA

XI. Acknowledgements

The report's authors gratefully acknowledge NSF, BBSRC and DFG for their contribution towards the travel costs of the attendees and meeting organization and Dr. Jan Witkowski and the staff of the Banbury Center for their professional hosting of the event.

Professor Mike Gale sadly died in July 2009. The report's authors are very grateful to have had the opportunity to work with him during the production of this vision document.

XII. References and Reports

Alston, Beddow and Pardey (2009), *Agricultural Research, Productivity, and Food Prices in the Long Run*. Science **325**:1209-10.

Burke MF, DB Lobell and L Guarino (2009) *Shifts in African crop climates by 2050, and the implications for crop improvement and general resource conservation*. Global Env Change **19**: 317-325

Campbell JE, Lobell DB, Genova RC and CB Field (2008) *The global potential of bioenergy on abandoned agricultural lands*. Env Sci Technol **42**: 5791-5794

Doos, BR (2002) *Population growth and loss of arable land*. Global Env. Change **12**: 303-311

Klein Goldewijk K, van Drecht G, Bouwman AF (2007) *Mapping contemporary global cropland and rangeland on a 5' resolution grid*. J Land Use Sci **2**:167–190

Scherr SJ and S Yadav (1996) *Land degradation in the developing world. Implications for food, agriculture and the environment to 2010*. Available online http://www.ifpri.org/sites/default/files/publications/vp14.pdf

2020 Vision for Biology (2008) The role of plants in addressing grant challenges in biology. Available online http://www.arabidopsis.org/portals/masc/workshop2020.pdf

European 2020 Vision for Plant Science (2009) Available online http://www.arabidopsis.org/portals/masc/2020_European_Vision.pdf

Royal Society (2009) Reaping the Benefits: *Science and the sustainable intensification of global agriculture*. Available online http://royalsociety.org/Reapingthebenefits/

Tackling Land Degradation and Desertification. (2002) GEF-IFAD Partnership. Available online http://www.ifad.org/events/wssd/gef/gef_ifad.htm

XIII. Discussion of Draft Document

The draft document was presented and discussed at the six international meetings and symposia in 2009.

XIV. Glossary

Apomixis:
Asexual reproduction, not including propagation from cuttings or leaves

Biofuel:
Fuel derived from renewable resources, especially plant biomass

Biome:
A major regional or global biotic community, such as a grassland or desert, characterized chiefly by the dominant forms of plant life and the prevailing climate

Bioproduct:
A material derived from a renewable resource, such as plastic made from plant oil

Curation:
The preservation and maintenance of a collection of data

Cytoskeletal:
Relating to the structural components within a cell's cytoplasm

Domestication:
The process in which an organism is adapted so that it can be cultivated for the benefit of humans

Germplasm:
A collection of genetic material/resources that comprises the physical basis of the inherited qualities of an organism. For plants, the germplasm may be stored as a seed collection.

Heterosis:
An increase in vigor or other improved qualities that often arises from crossbreeding; also known as hybrid vigor

Ideotype:
An idealized plant type with specific combinations of favorable characteristics

Metabolomic:
Relating to the study of the metabolome, the molecular intermediaries and products of metabolism within a cell

Metagenomic:
Relating to the study of the metagenome, the collective genetic material recovered from environmental samples

Multiscale modeling:
Models of systems that cross and link several spatial and/or temporal scales

'Omics:
The collection of fields of biology ending in "omics," including genomics, proteomics, and metabolomics

Pathogen:
A disease-causing organism, such as a virus or bacterium

Perenniality:
The quality of a plant in which its life cycle is two years or longer

Perturbagen:
Small molecules—peptides or fragments of proteins—that can be inserted into cells to create changes in phenotype

Phenotyping:
The determination of the physical characteristics of an organism

Photoperception:
The perception and reception of light

Phytopathology:
The area of plant research dealing with diseases

Phytosphere:
All environments relating to plants

Proteomic:
Relating to the proteome, the complete set of proteins expressed by a cell

Recombination:
The formation of new combinations of genes through natural processes, such as crossing over and independent assortment, or in the laboratory through the manipulation of genetic material

Self-incompatibility:
Any of several genetic mechanisms in plants that prevent self-fertilization

Spatio-temporal:
Relating to both space and time

Symbiont:
Any organism living in a state of symbiosis

Symbiosis:
Two dissimilar organisms living together; examples include mutualism and parasitism

Systems biology:
The study of the relationships between the various parts of the biological system with the aim of integrating this information to determine biological functions